**IN THE UNITED STATES
PATENT AND TRADEMARK OFFICE**

Appl. No.	: **09/351,086**
Applicant(s)	: **Nevenka Dimotrova**
Filed	: **09 Jul 1999**
TC/A.U.	: **2611**
Examiner	: **Bui, Kieu Oanh T**
Atty. Docket	: **PHA 23,716**

Title: **METHOD AND APPARATUS FOR LINKING A VIDEO SEGMENT TO ANOTHER SEGMENT OR INFORMATION SOURCE**

Mail Stop: **APPEAL BRIEF - PATENTS**
Commissioner for Patents
Alexandria, VA 22313-1450

APPEAL UNDER 37 CFR 41.37

Sir:

This is an appeal from the decision of the Examiner dated 24 August 2005, finally rejecting claims 1-25 of the subject application.

This paper includes (each beginning on a separate sheet):

1. **Appeal Brief, with appendices; and**
2. **Credit card authorization in the amount of $160, an appeal having been previously filed on 10 November 2004 in this case, and a fee of $340 having been paid.**

Adjustment date: 01/23/2006 TL0111
11/12/2004 MAHMED1 00000024 09351086
01 FC:1402 -340.00 OP

01/23/2006 TL0111 00000003 09351086
01 FC:1402 500.00 OP

APPEAL BRIEF

I. REAL PARTY IN INTEREST

The above-identified application is assigned, in its entirety, to **Philips Electronics North America**

II. RELATED APPEALS AND INTERFERENCES

Appellant is not aware of any co-pending appeal or interference that will directly affect, or be directly affected by, or have any bearing on, the Board's decision in the pending appeal.

III. STATUS OF CLAIMS

Claims 1-25 are pending in the application.

Claims 1-10 and 18-25 stand rejected by the Examiner under 35 U.S.C. 102(e).

Claims 11-17 stand rejected by the Examiner under 35 U.S.C. 103(a).

These rejected claims are the subject of this appeal.

IV. STATUS OF AMENDMENTS

No amendments were filed subsequent to the final rejection in the Office Action dated 24 October 2005. A reply to the final rejection was filed on 4 October 2005.

V. SUMMARY OF CLAIMED SUBJECT MATTER

The invention comprises a method and system for determining and accessing ancillary information regarding a feature in a video segment being displayed to a user (FIG. 1, page 5, lines 3-18). Of particular note, this on-line determination of an association allows a user to access information regarding a feature in the video, regardless of whether the original input video 12 includes links (hyperlinks) to this other information.

The determined association may be based on a semantic relationship, visual similarity, scene similarity, event similarity, and so on (page 6, line 18 – page 7, line 9; page 8, line 3 - page 9, line 21). When the user selects the object within the image, the associated source is accessed, and the associated information is displayed or stored for later viewing (page 7, lines 10-13; FIG. 6, page 14, lines 12-18). Alternatively, the determined information from the other source may be displayed automatically, using, for example, a picture-in-picture (PIP) presentation of available material (page 15, lines 4-10).

As claimed in independent claim 1, upon which claims 2-17 depend, the invention comprises a method for processing video, the method comprising:

determining an association between a first video segment including a particular feature and at least one additional information source also including that feature (page 2, lines 18-20; see also FIGs. 4-5); and

utilizing the association to display information from the additional information source (page 2, line 21 – page 3, line 2) based at least in part on a selection by a user of the feature in the first video segment while the video segment is displayed to the user (page 3, lines 10-14; FIG. 6, page 14, line 12 – page 17, line 5).

As claimed in claim 7, upon which claim 8 depends, the feature is a video feature extracted from at least one frame of the video segment (page 12, line 16 - page 13, line 10).

As claimed in dependent claim 9, upon which claims 10-13 depend, the feature is an audio feature extracted from at least one frame of the video segment (page 13, lines 17-18).

As claimed in dependent claim 10, the method includes combining an audio signal corresponding to the audio feature with an audio signal associated with the first video segment (page 15, lines 16-18).

As claimed in dependent claim 11, the method includes converting an audio signal corresponding to the audio feature into a textual format which is displayed with the first video segment (page 15, lines 16-18).

As claimed in dependent claim 12, the method includes separating at least a portion of the video segment into audio categories including one or more of single-voice speech, multiple-voice speech, music, silence and noise in order to extract the audio feature therefrom (page 13, lines 17-20).

As claimed in dependent claim 13, the audio feature comprises at least one of a music signature extraction, a speaker identification, and a transcript extraction (page 14, lines 3-4).

As claimed in dependent claim 14, upon which claim 15 depends, the feature is a textual feature extracted from at least one frame of the video segment (page 10, lines 17-19).

As claimed in dependent claim 15, the method includes displaying information corresponding to the textual information as an overlay on a display of the first video segment (page 15, lines 12-14).

As claimed in claim 16, the method includes determining the association based at least in part on at least one multi-dimensional feature vector extracted from a portion of the video segment using a feature extraction technique (page 10, lines 1-2).

As claimed in claim 17, the method includes determining the association based at least in part on at least one of a similarity measure and a clustering technique (page 10, lines 2-5).

As claimed in independent claim 18, the invention comprises an apparatus (FIG. 1) for processing video, the apparatus comprising:

a memory (16) for storing an association between a first video segment including a particular feature and at least one additional information source also including that feature (page 7, lines 4-9; page 10, lines 14-16; page 16, lines 12-19); and

a processor (15) coupled to the memory (16) and operative to utilize the association to direct the display of information from the additional information source based at least in part on a selection by a user of the feature in the first video segment while the video segment is displayed to the user (page 6, lines 3-12; page 12, lines 7-15; FIG. 6, page 14, lines 12-18).

As claimed in independent claim 19, the invention comprises an apparatus (FIG. 1) for processing video, the apparatus comprising:

a processor (15) operative

(i) to determine an association between a first video segment including a particular feature and at least one additional information source also including that feature (page 7, lines 4-9; page 16, lines 12-19; see also FIGs 4-5); and

(ii) to utilize the association to display information from the additional information source based at least in part on a selection by a user of the feature in the first video segment while the video segment is displayed to the user (page 6, lines 3-12; page 12, lines 7-15; FIG. 6, page 14, lines 12-18).

As claimed in independent claim 20, the invention comprises an article of manufacture comprising a machine-readable medium (page 19, lines 5-12) containing one or more software programs which when executed:

determine an association between a first video segment including a particular feature and at least one additional information source also including that feature (page 7, lines 4-9; page 16, lines 12-19; see also FIGs 4-6); and

utilize the association to display information from the additional information source based at least in part on a selection by a user of the feature in the first video segment while the video segment is displayed to the user (page 6, lines 3-12; page 12, lines 7-15; FIG. 6, page 14, lines 12-18).

As claimed in independent claim 21, the invention comprises a method for processing video, the method comprising:

determining from information in a portion of a first video segment an association between a particular feature of the first video segment and at least one additional information source also including that feature (page 2, lines 18-20; page 7, lines 4-9; page 10, lines 14-16; page 16, lines 12-19); and

utilizing the association to enable display of information from the additional information source when the first video segment is displayed (page 6, lines 3-12; page 12, lines 7-15; FIG. 6, page 15, lines 11-21).

As claimed in independent claim 22, the invention comprises an apparatus (FIG. 1) for processing video, the apparatus comprising:

a processor (15) operative

(i) to determine from information in a portion of a first video segment an association between a particular feature of the first video segment and at least one additional information source also including that feature (page 2, lines 18-20; page 7, lines 4-9; page 10, lines 14-16; page 16, lines 12-19); and

(ii) to utilize the association to enable display of information from the additional information source when the first video segment is displayed (page 6, lines 3-12; page 12, lines 7-15; FIG. 6, page 15, lines 11-21).

As claimed in independent claim 23, the invention comprises an article of manufacture comprising a machine-readable medium (page 19, lines 5-12) containing one or more software programs which when executed:

determine from information in a portion of a first video segment an association between a particular feature of the first video segment and at least one additional information source also including that feature (page 2, lines 18-20; page 7, lines 4-9; page 10, lines 14-16; page 16, lines 12-19); and

utilize the association to enable display of information from the additional information source when the video segment is displayed (page 6, lines 3-12; page 12, lines 7-15; FIG. 6, page 15, lines 11-21).

As claimed in independent claim 24, the invention comprises a method for processing video, the method comprising:

determining an association between (1) a portion of a video signal, the portion including a feature, and (2) at least one other information source also including the feature (page 7, lines 4-9; page 16, lines 12-19; see also FIGs 4-6); and

utilizing the association to enable display of information from the other information source (page 6, lines 3-12; page 12, lines 7-15) based at least in part on a selection by a user of the feature in the portion of the video signal while the video segment is displayed to the user (page 6, lines 3-12; page 12, lines 7-15; FIG. 6, page 14, lines 12-18).

As claimed in independent claim 25, the invention comprises a method for processing video, the method comprising:

determining from information in a portion of a video signal an association between a feature of the video signal and at least one other information source also including that feature (page 2, lines 18-20; page 7, lines 4-9; page 10, lines 14-16; page 16, lines 12-19); and

utilizing the association to enable display of information from the other information source when the video segment is displayed (page 6, lines 3-12; page 12, lines 7-15; FIG. 6, page 15, lines 11-21).

VI. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

Claims 1-10 and 18-25 stand rejected under 35 U.S.C. 102(e) over Hjelsvold et al. (USP 6,546,555, hereinafter Hjelsvold).

Claims 11-16 stand rejected under 35 U.S.C. 103(a) over Hjelsvold and Jain et al. (USP 6,463,444, hereinafter Jain).

Claim 17 stands rejected under 35 U.S.C. 103(a) over Hjelsvold.

VII. ARGUMENT

Claims 1-10 and 18-25 stand rejected under 35 U.S.C. 102(e) over Hjelsvold

MPEP 2131 states:

> "A claim is anticipated only if ***each and every element*** as set forth in the claim is found, either expressly or inherently described, in a single prior art reference." *Verdegaal Bros. v. Union Oil Co. of California*, 814 F.2d 628, 631, 2 USPQ2d 1051, 1053 (Fed. Cir. 1987). "The ***identical invention*** must be shown in as ***complete detail*** as is contained in the ... claim." *Richardson v. Suzuki Motor Co.*, 868 F.2d 1226, 1236, 9 USPQ2d 1913, 1920 (Fed. Cir. 1989).

Claims 1-10 and 19-25

Each of independent claims 1 and 19-25 include the element of determining an association between a first video segment including a particular feature and at least one additional information source also including that feature.

Hjelsvold fails to teach determining an association between a first video segment including a particular feature and at least one additional information source also including that feature.

Hjelsvold teaches a system for selling video information over a communications network. Hjelsvold teaches filtering existing hypervideo sequences to create a customized video sequence based on the 'price class' that a user selects. As illustrated in Hjelsvold's FIG. 1, and detailed at column 4, line 65 through column 5, line 13, a meta-data database 12 contains the AIUs (Anchorable Information Units) and hyperlinks that connect these units. That is, the database 12 contains all of the defined hyperlinks that are available for the video segments (NSS – Narrative Sequences of Scenes) contained in the media data 11. A filtering process 18 controls a media file generator 17 to select/filter the available AIUs based on information received from a payment client 24. The selected/filtered AIUs control the access and sequencing of segments from the available media data 11. As illustrated in FIGs. 17 and 18, for example, the sequence provided to a user is dependent upon whether the user is a 'premium customer', an 'affiliated customer', or a 'standard customer'.

Hjelsvold teaches selecting from among a set of predefined hyperlinks/associations; if an AIU does not already exist to provide a link to additional information, Hjelsvold's system does not analyze the video segment to determine an association to provide such a link. Hjelsvold teaches using existing links, and does not teach determining associations between video segments and other information sources, as specifically taught and claimed by the applicant.

Because Hjelsvold fails to teach determining an association between a first video segment including a particular feature and at least one additional information source also including that feature, the applicant respectfully maintains that the rejection of claims 1-10 and 19-25 under 35 U.S.C. 102(e) over Hjelsvold is unfounded.

Claims 7-8

As claimed in claim 7, upon which claim 8 depends, the feature is a video feature extracted from at least one frame of the video segment.

Hjelsvold fails to teach extracting a video feature from a frame.

The Office action asserts that Hjelsvold teaches extracting a video feature from a frame in FIG. 17 and at column 12, lines 1-33. The applicant respectfully disagrees with this assertion. FIG. 17 illustrates the motion of an object having an associated AIU on a display, and shows a "next frame" that will be displayed. Hjelsvold does not teach that the next frame includes a feature that is "extracted" from a frame of the sequence, as asserted in the Office action.

Because Hjelsvold fails to teach extracting a video feature from a frame, the applicant respectfully maintains that the rejection of claims 7-8 under 35 U.S.C. 102(e) over Hjelsvold is unfounded.

Claims 9-10

As claimed in claim 9, upon which claims 10-13 depend, the feature is an audio feature extracted from at least one frame of the video segment.

Hjelsvold fails to teach extracting an audio feature from a frame.

The Office action asserts that Hjelsvold teaches extracting an audio feature from a frame at column 4, lines 51-64. The applicant respectfully disagrees with this assertion. The cited text merely recites that the media data includes audio information. Hjelsvold does not teach that any of this audio information is "extracted" from a frame of the sequence to form the feature that is used to determine an association, as asserted in the Office action.

Because Hjelsvold fails to teach extracting an audio feature from a frame, the applicant respectfully maintains that the rejection of claims 9-10 under 35 U.S.C. 102(e) over Hjelsvold is unfounded.

Claim 10

As claimed in dependent claim 10, the method includes combining an audio signal corresponding to the audio feature with an audio signal associated with the first video segment.

Hjelsvold does not teach combining audio signals. Hjelsvold teaches that the frames identified by the AIUs are rendered sequentially, and does not teach mixing or merging the audio among the selected segments.

Because Hjelsvold fails to teach combining an audio signal corresponding to the audio feature with an audio signal associated with the first video segment, the applicant respectfully maintains that the rejection of claim 10 under 35 U.S.C. 102(e) over Hjelsvold is unfounded.

Claim 18

Claim 18 claims an apparatus that includes a processor that directs the display of information from an additional information source based at least in part on a selection by a user of a feature in the first video segment while the video segment is displayed to the user.

Hjelsvold fails to teach controlling a display based on a selection by a user of the feature in the first video segment while the video segment is displayed to the user. The Office action asserts that FIGs. 16-18 and the associated text teach the

selection of a feature in the displayed video segment. The applicant respectfully disagrees with this assertion, and note that the cited text does not address the selection of a feature in the displayed video.

Because Hjelsvold fails to teach selection by a user of a feature in the first video segment while the video segment is displayed to the user, the applicant respectfully maintains that the rejection of claim 18 under 35 U.S.C. 102(e) over Hjelsvold is unfounded.

Claims 11-16 stand rejected under 35 U.S.C. 103(a) over Hjelsvold and Jain

MPEP 2142 states:

> "To establish a *prima facie* case of obviousness ... the prior art reference (or references when combined) ***must teach or suggest all the claim limitations***... If the examiner does not produce a *prima facie* case, the applicant is under no obligation to submit evidence of nonobviousness."

Claims 11-16

Each of claims 11-16 is dependent upon claim 1, which includes the limitation of determining an association between a first video segment including a particular feature and at least one additional information source also including that feature. The Office action relies upon Hjelsvold for teaching all of the elements of claim 1.

As detailed above, Hjelsvold fails to teach determining an association between a first video segment including a particular feature and at least one additional information source also including that feature.

Because Hjelsvold fails to teach determining an association between a first video segment including a particular feature and at least one additional information source also including that feature, the applicant respectfully maintains that the rejection of claims 11-16 under 35 U.S.C. 103(a), which relies upon Hjelsvold for this teaching, is unfounded, per MPEP 2142.

Claims 14-15

As claimed in dependent claim 14, upon which claim 15 depends, the feature is a textual feature extracted from at least one frame of the video segment.

Hjelsvold and Jain fail to teach or suggest extracting a textual feature from a frame of the video segment.

The Office action asserts that Hjelsvold teaches that the feature is a textual feature that is extracted from the video frame, but provides no basis for this assertion.

Because Hjelsvold fails to teach a textual feature extracted from at least one frame of the video segment, the applicant respectfully maintains that the rejection of claims 14-15 under 35 U.S.C. 103(a), which relies upon Hjelsvold for this teaching, is unfounded, per MPEP 2142.

Claim 16

As claimed in claim 16, the method includes determining the association based at least in part on at least one multi-dimensional feature vector extracted from a portion of the video segment using a feature extraction technique.

Hjelsvold and Jain fail to teach or suggest determining an association based at least in part on at least one multi-dimensional feature vector extracted from a portion of the video segment using a feature extraction technique.

The Office action asserts that Hjelsvold provides this teaching at FIG. 14 and column 12, lines 20-46. The applicant respectfully disagrees with this assertion. FIG. 14 illustrates a streaming server providing a video stream to a client. Column 12, lines 20-46 detail how a merchant selects promotion material to be added to the streamed video. Neither of these cited references address feature extraction from a portion of the video segment.

Because Hjelsvold fails to teach determining an association based at least in part on at least one multi-dimensional feature vector extracted from a portion of the video segment using a feature extraction technique, the applicant respectfully maintains that the rejection of claim 16 under 35 U.S.C. 103(a), which relies upon Hjelsvold for this teaching, is unfounded, per MPEP 2142.

Claim 17 stands rejected under 35 U.S.C. 103(a) over Hjelsvold

Claim 17

Claim 17 is dependent upon claim 1, which includes the limitation of determining an association between a first video segment including a particular feature and at least one additional information source also including that feature.

As detailed above, Hjelsvold fails to teach determining an association between a first video segment including a particular feature and at least one additional information source also including that feature.

Because Hjelsvold fails to teach determining an association between a first video segment including a particular feature and at least one additional information source also including that feature, the applicant respectfully maintains that the rejection of claim 17 under 35 U.S.C. 103(a) over Hjelsvold is unfounded, per MPEP 2142.

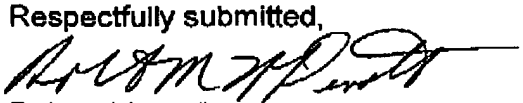
CONCLUSIONS

Because Hjelsvold fails to teach determining an association between a first video segment including a particular feature and at least one additional information source also including that feature, the applicant respectfully requests that the Examiner's rejection of claims 1-10 and 19-25 under 35 U.S.C. 102(e), and claims 11-17 under 35 U.S.C. 103(a) be reversed by the Board, and the claims be allowed to pass to issue.

Because Hjelsvold fails to teach selection by a user of a feature in the first video segment while the video segment is displayed to the user, the applicant respectfully requests that the Examiner's rejection of claim 18 under 35 U.S.C. 102(e) be reversed by the Board, and the claims be allowed to pass to issue.

In the alternative, because Hjelsvold also fails to teach the additional limitations of claims 7-10 and 14-16, as detailed above, the applicant respectfully requests that the Examiner's rejection of each of these claims be reversed by the Board, and the claims be allowed to pass to issue.

Respectfully submitted,



Robert M. McDermott, Attorney
Registration Number 41,508
804-493-0707

CLAIMS APPENDIX

1. A method for processing video, the method comprising:

determining an association between a first video segment including a particular feature and at least one additional information source also including that feature; and

utilizing the association to display information from the additional information source based at least in part on a selection by a user of the feature in the first video segment while the video segment is displayed to the user.

2. The method of claim 1 wherein

determining the association further includes

retrieving the association from a memory.

3. The method of claim 1 wherein

determining the association further includes

determining the association from information in a portion of the video segment.

4. The method of claim 1 wherein

the additional information source comprises

an additional video segment also including the feature.

5. The method of claim 4 wherein

utilizing the association includes

switching from display of the first video segment to display of the additional video segment also including the feature.

6. The method of claim 4 wherein

utilizing the association includes

displaying the additional video segment at least in part in a separate portion of a display which also includes at least a portion of the first video segment.

7. The method of claim 1 wherein

the feature is a video feature extracted from at least one frame of the video segment.

8. The method of claim 7 wherein

the video feature comprises at least one of

a frame characterization,

a face identification,

a scene identification,

an event identification, and

an object identification.

9. The method of claim 1 wherein

the feature is an audio feature extracted from at least one frame of the video segment.

10. The method of claim 9 wherein

utilizing the association includes

combining an audio signal corresponding to the audio feature with an audio signal associated with the first video segment.

11. The method of claim 9 wherein

utilizing the association includes

converting an audio signal corresponding to the audio feature into a textual format which is displayed with the first video segment.

12. The method of claim 9 further including

separating at least a portion of the video segment into audio categories including one or more of

single-voice speech,

multiple-voice speech, music,

silence and

noise

in order to extract the audio feature therefrom.

13. The method of claim 9 wherein

the audio feature comprises at least one of

a music signature extraction,

a speaker identification, and

a transcript extraction.

14. The method of claim 1 wherein

the feature is a textual feature extracted from at least one frame of the video segment.

15. The method of claim 14 wherein

utilizing the association includes

displaying information corresponding to the textual information as an overlay on a display of the first video segment.

16. The method of claim 1 wherein

determining the association further includes

determining the association based at least in part on at least one multi-dimensional feature vector extracted from a portion of the video segment using a feature extraction technique.

17. The method of claim 1 wherein

determining the association further includes

determining the association based at least in part on at least one of a similarity measure and a clustering technique.

18. An apparatus for processing video, the apparatus comprising:

a memory for storing an association between a first video segment including a particular feature and at least one additional information source also including that feature; and

a processor coupled to the memory and operative to utilize the association to direct the display of information from the additional information source based at least in part on a selection by a user of the feature in the first video segment while the video segment is displayed to the user.

19. An apparatus for processing video, the apparatus comprising:

a processor operative

(i) to determine an association between a first video segment including a particular feature and at least one additional information source also including that feature; and

(ii) to utilize the association to display information from the additional information source based at least in part on a selection by a user of the feature in the first video segment while the video segment is displayed to the user.

20. An article of manufacture comprising a machine-readable medium containing one or more software programs which when executed:

determine an association between a first video segment including a particular feature and at least one additional information source also including that feature; and

utilize the association to display information from the additional information source based at least in part on a selection by a user of the feature in the first video segment while the video segment is displayed to the user.

21. A method for processing video, the method comprising:

determining from information in a portion of a first video segment an association between a particular feature of the first video segment and at least one additional information source also including that feature; and

utilizing the association to enable display of information from the additional information source when the first video segment is displayed.

22. An apparatus for processing video, the apparatus comprising:

a processor operative

(i) to determine from information in a portion of a first video segment an association between a particular feature of the first video segment and at least one additional information source also including that feature; and

(ii) to utilize the association to enable display of information from the additional information source when the first video segment is displayed.

23. An article of manufacture comprising a machine-readable medium containing one or more software programs which when executed:

determine from information in a portion of a first video segment an association between a particular feature of the first video segment and at least one additional information source also including that feature; and

utilize the association to enable display of information from the additional information source when the video segment is displayed.

24. A method for processing video, the method comprising:

determining an association between (1) a portion of a video signal, the portion including a feature, and (2) at least one other information source also including the feature; and

utilizing the association to enable display of information from the other information source based at least in part on a selection by a user of the feature in the portion of the video signal while the video segment is displayed to the user.

25. A method for processing video, the method comprising:

determining from information in a portion of a video signal an association between a feature of the video signal and at least one other information source also including that feature; and

utilizing the association to enable display of information from the other information source when the video segment is displayed.

EVIDENCE APPENDIX

No evidence has been submitted that is relied upon by the appellant in this appeal.

RELATED PROCEEDINGS APPENDIX

Appellant is not aware of any co-pending appeal or interference which will directly affect or be directly affected by or have any bearing on the Board's decision in the pending appeal.